

WSN-Driven Posture Recognition and Correction Towards Basketball Exercise

Xiangyang Cai, Zhengzhou University of Industry Technology, China*

ABSTRACT

In order to enhance the daily training for basketball, this paper establishes a human posture estimation framework by using monocular camera and wireless sensor network. First, the daily basketball training images are collected by monocular camera and transmitted through wireless sensor network. Second, the collected images are processed by an observation and reasoning model which is based on component and graph reasoning. The basketball player's posture is depicted by the rotation invariant features of edge field. The extracted features are used to learn a boosting classifier as the observation model. The experimental results show that the posture recognition rate can achieve more than 88% for basketball player's action.

KEYWORDS

Boosting Classifier, Observation Model, Posture Estimation, Sport Action Recognition

1. INTRODUCTION

Human posture estimation and recognition (Jalal et al. 2020) can be applied to human activity analysis (Murahari et al. 2018; Batool et al. 2019), human-computer interaction (Xu et al. 2019; Grosse-Puppenthal et al. 2017) and visual surveillance (Kumaran et al. 2019; Sharma et al. 2019), which is a hot topic in the field of computer vision. In the process of basketball training and competition, coaches need to make special training plans according to the individual situation of different players. In order to improve the basketball skills of athletes, the traditional training method is that coaches make training plan according to their own training theory and experience by combining with the skill level of basketball players. This training mode is highly subjective and coaches need to spend a lot of time to analyze the athletes' posture. It is difficult to objectively evaluate the athletes' training effect. The core of modern sports training is accuracy and efficiency. If the coach can accurately control the athletes' movement posture training, the effect can be greatly improved. It has become a new research direction to collect and analyze the posture data of basketball players and accurately identify the movement posture by using wireless sensor network (WSN) (Tomić et al. 2017; Wang et al. 2017) and computer vision, which is of great significance to improve the coaches' training plan and the training effect of athletes.

DOI: 10.4018/IJISMD.300777

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

Basketball posture recognition (Fan et al. 2021; Ren et al. 2021) is a kind of human posture recognition. At present, there are two main methods of human posture recognition: posture recognition based on inertial sensor (Wang et al. 2019; Zhang et al. 2018) and posture recognition based on image acquisition (Chevtchenko et al. 2018; Oudah et al. 2020). According to the number of image acquisition devices, posture recognition based on image acquisition can be further classified as monocular video recognition (Cai et al. 2019) and multicular video recognition (Mircoli et al. 2018). The general idea of posture recognition based on image acquisition first adopts the camera to capture the athletes' images or videos, then extracts the hidden motion features, and finally designs a classifier to recognize the athletes' motion posture. The multicular video based posture recognition has high maturity and high accuracy. However, the monocular video based posture recognition needs to be further researched at this stage. First, the human body is composed of multiple components, the motion is very complex, and there lacks 3D information in monocular pictures and videos. The change of human 2D posture cannot be simply described by a unified model. Second, the background and light in the image or video have great changes in different scenes. The clothing of the human body itself do not keep the same as well. These changes result in dramatic variation in the appearance of the human body, which is difficult to be described by a unified model.

The current work about posture recognition can be classified as whole based method (Bui et al. 2018) and the component-based method (Lahiani et al. 2017). The whole based method adopts the whole appearance or segmentation results of human body in the image to estimate the posture. The holistic approach can be divided into regression based approach (Li et al. 2019) and case based approach (Qi et al. 2021). The method based on regression (Li et al. 2020) regards the posture estimation problem as a direct mapping from human appearance to human posture, and learns the regression function from image features or segmentation results to describe human posture according to a large number of samples. The case based method adopts a representative sample set to represent the posture space, and encodes these samples as image features. When estimating the posture, the image features of the future posture are extracted to compare with the features of the existing postures in the posture space. The most similar posture in the posture space is just the predicted result. The whole based method is usually efficient, but difficult to cover the whole posture space with limited samples due to the complexity of human posture changes. The whole based method is only suitable for the estimation of specific posture.

In the component-based method, the human body is divided into several interrelated components and represented by a graph model. The human posture is optimized by graph reasoning method. There are three key points in the component-based method: graph model, component observation model and optimization algorithm. The graph model is used to represent the constraint relationship between components. The tree model is the most commonly used graph model. It is defined according to the connection relationship of components, which is the most intuitive. In order to solve the self-occlusion in the posture estimation and recognition, the constraint between non-adjacent components is introduced in tree model. Another way is to induce symmetry constraints between symmetrical components to enhance the tree model for posture estimation.

The optimization algorithm estimates the final posture according to the graph model and observation model. Belief propagation (BP) (Friston et al. 2017) is a commonly used graph model optimization algorithm. However, the dimensions of component state vector are very high in the problem of human posture estimation. It is not realistic to use BP algorithm directly. According to the characteristics of tree model, a pictorial structure algorithm is proposed to speed up the process of graph optimization. However, the pictorial structure algorithm limits the potential function between components. In order to solve this issue, an adaptive potential function is introduced into pictorial structure algorithm.

Aiming at the problem of human posture estimation in monocular images and videos for basketball player posture estimation and recognition, this paper adopts the component-based method to divide the human body into several parts and utilizes a tree model to represent the human body. The human

body is represented as a rotation invariant edge force field and the associated features are used to learn a classifier which is used to predict future player's posture. The rotation invariant edge force field is expressive and easy to be calculate, which can be used in the observation model for the components quickly and accurately. In addition, this paper adopts the idea of particle filter to propose an optimization algorithm based on particle sampling and BP algorithm. The algorithm can utilize iterative search to estimate the final attitude. The player images are collected through monocular camera and transmitted to processing center through wireless sensor network. An illustration of the proposed posture estimation framework is shown in Fig. 1.

The rest of this article is organized as follows. The problem description is provided in the Section 2. The component detection based posture estimation is provided in the Section 3. The experimental evaluation is provided in the Section 4. The section is the Conclusion.

2. PROBLEM DESCRIPTION OF THE POSTURE ESTIMATION AND RECOGNITION

The human body is divided into 10 interrelated parts: head, trunk, left and right upper arms, left and right forearms, left and right thighs and left and right calves, which are represented as a tree Markov network. An illustration is shown in Figure 1.

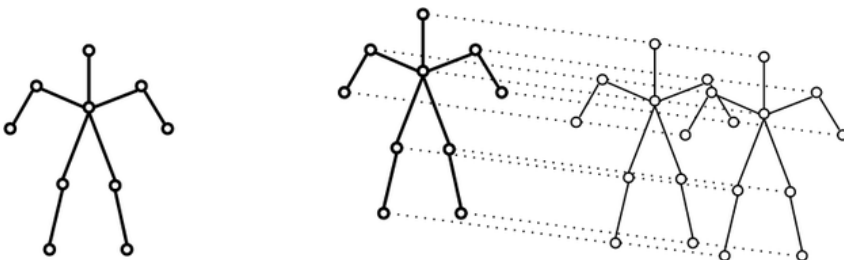
In Figure 2, each node represents a part of the human body and the associated status is represented as p_i . Each part has its own observation model, denoted as $P(I_i | p_i)$. The geometric constraints between adjacent parts are denoted as $P(p_i, p_j)$. The posture of human body is represented as the set of posture status of human components, denoted as $S = \{p_1, \dots, p_n\}$. The posture status of the i^{th} component is determined by location $l(x_i, x_j)$, size s_i , and direction θ_i . Then, the associated posture status is represented as $p_i = \{l_i, s_i, \theta_i\}$.

For a certain image appearance I , the associated human posture estimation and recognition can be formalized as the following optimization problem

Figure 1. The architecture of posture estimation by using wireless sensor network



Figure 2. The illustration of posture in Markov Hidden Model. In (a), human body is depicted by a tree structure. In (b), the posture is represented by a Markov network



$$\begin{aligned}
 L &= \operatorname{argmax} P(L|I) \\
 &= \operatorname{argmax} P(I|L) P(L) \\
 &= \prod_i P(I_i|p_i) \prod_{(i,j) \in E} P(p_i, p_j)
 \end{aligned} \tag{1}$$

In Equation (1), $P(I_i|p_i)$ is the observation model of each component, which will be described in the following part; $P(p_i, p_j)$ represents the geometric constraints between adjacent components. It is assumed that the constraints of position, size and direction are independent of each other. The geometric constraint is represented as follows:

$$\begin{aligned}
 P(p_i, p_j) &= P(l_i, l_j) P(s_i, s_j) P(\theta_i, \theta_j) \\
 &\propto N(l_i - l_j; 0, \sigma_l) N(s_i - s_j; 0, \sigma_s) e^{k \cos(\theta_i - \theta_j - \mu_{i,j})}
 \end{aligned} \tag{2}$$

In Equation (2), $N(l_i - l_j; 0, \sigma_l)$ and $N(s_i - s_j; 0, \sigma_s)$ represent the distributions of location and size. The distributions follow the normal distribution with mean 0 and variance σ_l for location and variance σ_s for size. These parameters can be automatically learnt.

In a video, the components in the human body in each frame can form a sequential Markov network, which is illustrated in Figure 2. The posture in current frame L_t can be obtained by current image I_t and the posture in previous frame L_{t-1} , which can be written as follows:

$$\begin{aligned}
 L_t &= \operatorname{argmin} \prod_i P(I_{t,i} | p_{t,i}) \prod_{(i,j) \in E} P(p_{t,i}, p_{t,j}) \\
 &\quad \times \prod_i P(p_{t,i} | p_{t-1,i})
 \end{aligned} \tag{3}$$

In Equation (3), $P(p_{t,i} | p_{t-1,i})$ is the motion model of the component i , which is written as follows

$$\begin{aligned}
 P(p_{t,i} | p_{t-1,i}) &= N(l_{t,i} - l_{t-1,i}, 0, \sigma_{l,i}) \\
 &\quad \times N(s_{t,i} - s_{t-1,i}, 0, \sigma_{s,i}) \\
 &\quad \times e^{k_i \cos(\theta_{t,i} - \theta_{t-1,i} - \mu_i)}
 \end{aligned} \tag{4}$$

The parameters $\sigma_{l,i}$, $\sigma_{s,i}$, k_i and μ_i can be obtained by learning.

3. COMPONENT DETECTION FOR POSTURE ESTIMATION AND RECOGNITION

In the field of computer vision and image processing, the features based on gradient are the most commonly used to represent image. The histogram of gradient (HOG) is one of the most commonly used gradient features. HOG features are widely used in pedestrian detection and human posture estimation. However, the calculation of HOG features needs a lot of preprocessing work on the image and is sensitive to the direction and size of the object. It is only suitable for fast detection of objects

with roughly determined direction in the image, such as pedestrians. However, the size and direction of human body components in the image change dramatically, which consumes a lot of time and space to calculate HOG features.

Another gradient based feature is Edgelet, which counts the edge response on a specific curve segment. Edgelet feature is easy to calculate. However, it only considers local gradient feature and is vulnerable to noises.

In human posture estimation, due to the influence of human motion posture and clothing, the components of human body do not only have variable direction and size, but also have rich texture. Therefore, the features must be stable rotation invariant and easy to calculate.

By referring to the basic physical knowledge, the gradient image can be regarded as a field which is full of “charged” particles. In the field, each pixel q_i can be regarded as a “charged” particle, which is termed as gradient particle. The associated “charged” quantity is defined as the gradient intensity $|d_{x_i}|$. There exists interaction force between each pair of the gradient particles, which is similar to the electronic force. The interaction force between gradient particle q_i and q_j is defined as follows:

$$f(q_i, q_j) = \frac{d_{q_i} \cdot d_{q_j}}{r^2} \quad (5)$$

In Equation (5), $r = q_i - q_j$ is the vector from q_j to q_i . Equation (5) represents that the force between two gradient particles is directly proportional to their gradient amplitude, directly proportional to the cosine of gradient direction difference, and inversely proportional to the square of distance. In the gradient image I , a gradient particle q_i is placed at position l , and the force acting on it is written as follows:

$$f(q_i, I) = \sum_{q_j \in \Gamma(q_i)} \frac{d_{q_i} \cdot d_{q_j}}{r_{q_j}^3} r_{q_j} \quad (6)$$

In Equation (6), $\Gamma(q_i) \in I$ represents the neighborhood of q_i in I and r_{q_j} represents the vector from q_j to q_i .

Obviously, $f(q_i, I)$ generally points to the direction of the strongest edge response near q_i in I . The orthogonal decomposition of $f(q_i, I)$ is written as follows:

$$\begin{aligned} f(q_i, I) &= \left(f(q_i, I) \cdot e_x, f(q_i, I) \cdot e_y \right)^T \\ &= \begin{pmatrix} d_{q_i} \cdot \sum_{q_j \in \Gamma(q_i)} \frac{r_{q_j} \cdot e_x}{r_{q_j}^3} \\ d_{q_i} \cdot \sum_{q_j \in \Gamma(q_i)} \frac{r_{q_j} \cdot e_y}{r_{q_j}^3} \end{pmatrix} \\ &= \begin{pmatrix} d_{q_i} \cdot F_x(q_i) \\ d_{q_i} \cdot F_y(q_i) \end{pmatrix} \\ &= d_{q_i} \cdot F(q_i) \end{aligned} \quad (7)$$

In Equation (7), e_x and e_y represents the positive unit vector along x-axis and y-axis, respectively; $F(q_i) = (F_x(q_i)^T, F_y(q_i)^T)$ is a 2×2 matrix which is termed as edge force field. The edge force field is only related with the location of q_i .

In fact, the force $f(q_i, I)$ can be regarded as the weighted sum of the projections of the image gradients in the direction of d_{q_i} in the neighborhood of position q_i . An illustration is shown in Figure 2.

In Figure 3, the features of edge force field are defined as a curve segment C which composes of several gradient particles. The eigenvalue is defined as the projection of the sum of the forces on all particles in the edge force field on the unit vector u , which is written as follows:

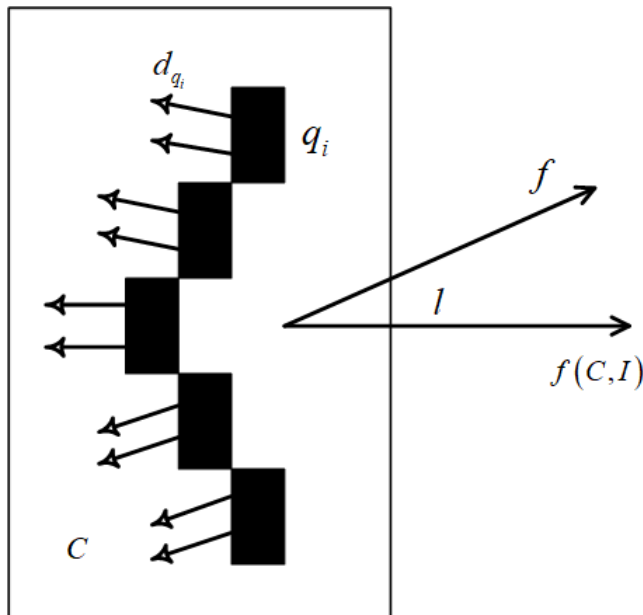
$$f(C, I) = \sum_{q \in C} F(q, I) \cdot u = \sum_{q \in C} d_q^T F(q) \cdot u \quad (8)$$

Because the eigenvalue is the inner product of the vector, the calculation of the eigenvalue is rotation invariant which means the features are not sensitive to the direction the posture. The edge force field $F(q, I)$ is independent of the specific features, and can be pre-calculated by convolution according to the original image, which is written as follows:

$$F_{i,j} = I \otimes D_i \otimes T_j, \quad i, j \in \{x, y\} \quad (9)$$

In Equation (9), D_i is convolution kernel for calculating image gradient, while T_j is convolution kernel whose intensity decreases with the radius. According to the characteristics of the edge force

Figure 3. The illustration of edge force field. The eigenvalue $f(S, I)$ is defined as the sum of the forces on all particles in the edge force field



field, It can simply design a weak classifier based on piecewise linear function. The classifier divides the range of corresponding eigenvalues into several disjoint subspaces $\{(t_i, t_{i+1}]\}$ with equal interval. All samples in the same interval are mapped into a fixed prediction result which is written as follows:

$$h(C, I) = v_i, \quad t_i < f(C, I) \leq t_{i+1} \quad (10)$$

The weak classifiers based on edge force field can be assembled by AdaBoost algorithm which can be regarded as the weighted sum of a series of weak classifiers. In classical AdaBoost learning algorithm, all features are enumerated in each iteration. The best weak classifier is selected according to the current sample weights. However, when edge force field features are used, the number of features in a sub-window of image is huge. The number of features of edge force field reaches near 10^3 for a sub-window with 30×30 . The structure of edge force field can be further divided into different features according to different projected directions. The exhaustive search method cannot be used to select weak classifiers directly. This paper adopts heuristic search method for weak classifier selection. The heuristic search for weak classifier selection is summarized as follows:

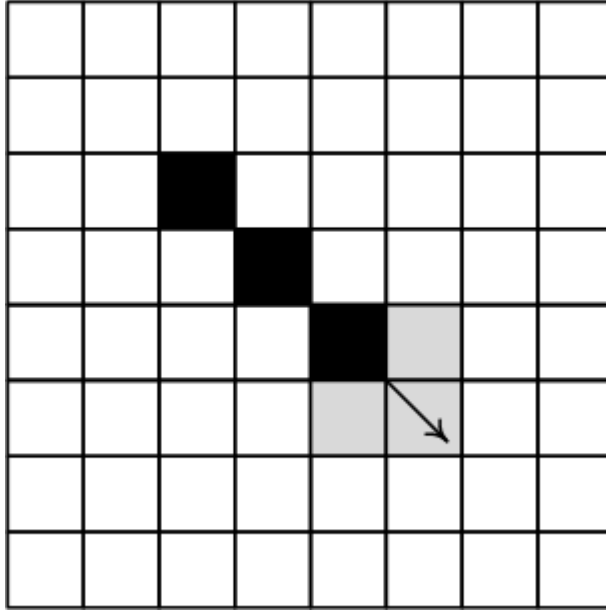
The Algorithm 1 first initializes the feature list as all the features with only one gradient particle, then selects the features with the minimum classification loss in the feature list in each iteration and generates a series of new features to add into the feature list according to the extended rules, lastly selects the features with the minimum classification loss as the features of the weak classifier after several iterations. A gradient particle is added in each expansion. In order to ensure that the same features are not generated in the expansion process, and to ensure the smoothness of the curve, the position of each gradient particle is limited. The constraint is illustrated in Figure 4.

In Figure 4, given the position q of gradient particle, it still needs to determine the projection direction o and the gradient of particle d_q to ensure that the training samples can be separated as

Algorithm 1 (weak classifier selection)

Input: training set $\{I_i\}$, sample weight set $\{D_i\}$ Initialization: The feature list is initialized as $FL_1\{C_1, \dots, C_n\}$ to calculate classification loss. The C_i is the feature set that is only from one gradient particle. The feature list after selection is denoted as $FL_2 = \emptyset$.
Weak classifier selection: For $i=1$ to n Step 1: Select the features with the least classification loss from feature list FL which is written as follows $C' = \underset{C \in FL}{\operatorname{argmin}} O(C)$ Here, $O(C) = \sum_i D_i e^{-h(C, I_i)}$, $FL_1 = FL_1 - \{C'\}$, and $FL_2 = FL_2 \cup \{C'\}$. Step 2: According to the rule extension C', it calculates the classification loss and adds the associated features in FL_2; End for
Output: Select the features with least classification loss from FL_1 and FL_2.

Figure 4. The limitation of position of gradient particle. The scale squares are the constraint directions



far as possible. Since d_q is not linear with the direction u , it needs to update the direction and gradient alternately as follows:

First, given direction u , the gradient is updated. For a certain direction, Equation (8) is converted as follows:

$$\begin{aligned} f(C, I) &= \sum_{q \in C} d_q^T F(q) u \\ &= (d_{q_1}^T, \dots, d_{q_n}^T) \cdot (F(q_1)u, \dots, F(q_n)u) \end{aligned} \quad (11)$$

Second, given gradient d_q , the direction is updated.

Belief propagation (BP) algorithm is a common graph reasoning algorithm, which estimates the marginal probability distribution of each node by iteratively propagating information between adjacent nodes in the graph. For Equation (1), the information propagated from node p_i to node p_j in each iteration is written as follows:

$$m_{p_i, p_j}^n = \alpha \int_{l_{p_i}} P(l_{p_i}, l_{p_j}) P(l_{p_i} | I) \prod_{u \in \Gamma(p_j)} m_{u, p_i}^n(l_{p_i}) dl_{p_i} \quad (12)$$

The marginal probability of each node can be estimated according to the observation model of the node and the information of adjacent nodes, which is written as follows:

$$\hat{P}(l_{p_i} | I) = \beta P(I | l_{p_i}) \prod_{u \in \Gamma(p_i)} m_{u, p_i}^n(l_{p_i}) \quad (13)$$

In Equation (12) and (13), α and β are normalization factors.

For the tree graph, BP algorithm can converge to the real marginal distribution. It is very time-consuming to calculate the integral in Equation (12) in the whole state space since the observation model of each node can only be measured discretely and the dimension of node's state vector dimension is high. In order to solve this issue, this paper adopts a particle sampling based belief propagation algorithm. For each node in the Markov network in Figure 2, this paper adopts a group of weighted particles $\{(p_i^0, \pi_i^0), \dots, (p_i^n, \pi_i^n)\}$ to simulate the distribution of node i . The maximum posteriori probability of the problem is obtained by continuous resampling. The procedure is summarized in the following algorithm.

In essence, the Algorithm 2 is a random variable step search algorithm.

Basketball players' posture recognition is to construct a classifier which can recognize the players' posture according to the data features that are extracted from the data collected by sensors. The extracted posture features are input into the classifier. Then, the classifier outputs a specific basketball action.

4. EXPERIMENTS AND SIMULATIONS

In this section, we collect eight postures from 12 basketball players. The postures include no action, walking, running, dribbling, jumping, shooting, passing and catching. The basketball players repeat each action 90 times. Then, we can obtain 1,080 samples for each action and 8,640 samples in total. All sample are denoted manually. In the process of motion posture data collection, the players complete the prescribed basketball action according to the preset posture and his usual exercise habits. Any basketball posture is composed of 10 interrelated parts: head, trunk, left and right upper arms, left and right forearms, left and right thighs and left and right calves. We test the component detection result and posture detection result on collected data, respectively. The collected samples are split as training set and test set. The size of training set is from 10% to 50% step by 10% of the all samples, while the rest samples are used as test set.

For the component detection, the features are extracted by Edge field. In addition, we also compare Edge field with histogram of gradient (HOG), and scale-invariant feature transform (SIFT). The curve between the detection rate and the size of training set is reported in Figure 4.

From Figure 5, it can be found that the component detection rate increases with the size of training set. The component detection rate is from 87.21% to 90.04% for Edge filed, from 85.30% to 88.13% for HOG features, and from 84.65% to 87.93% for SIFT features. When the size of training set reaches 50% of the collected data, the component detection rate reaches 90.04% for Edge field, 88.13% for HOG features, and 87.93% for SIFT features, respectively. It can be found that the Edge field performs better than HOG features and SIFT features.

Algorithm 2 (reasoning model)

Initialization:

The particles of each component are initialized according to the component detection result or the previous frame information

Step 1: the posterior marginal distribution of each component on particle set is calculated by BP algorithm and the weight is updated by using the following equation:

$$\pi_i^h = P(l_i^h | I) \quad (14)$$

Step 2: the particles of each component are resampled according to the weight;

Step 3: the particle with maximum posterior marginal distribution is adopted as final status. If the difference between this iteration and previous iteration is less than a threshold, the algorithm terminates; otherwise, the algorithm goes to the next iteration.

Furthermore, we compare the ROC curve for three feature extraction method when 50% samples are used as training set. The experimental result of ROC curve is reported in Figure 5.

From Figure 6, it can be found that the Edge field can achieve better ROC result than HOG and SIFT. The related AUC value reaches 0.8782, 0.8794, and 0.8887 for HOG features, SIFT features, and Edge field features, respectively.

After obtaining component detector, the human posture can be estimated by using particle sampling induced confidence propagation algorithm. For static image or the first frame of the

Figure 5. The curve between the detection rate and the size of training set

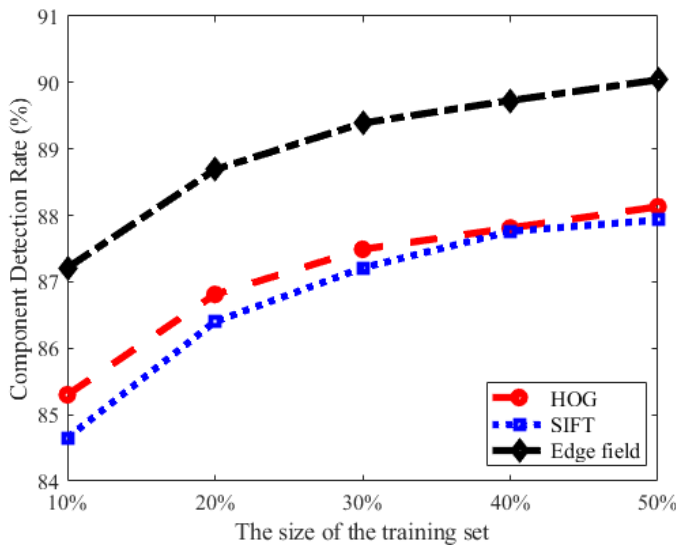
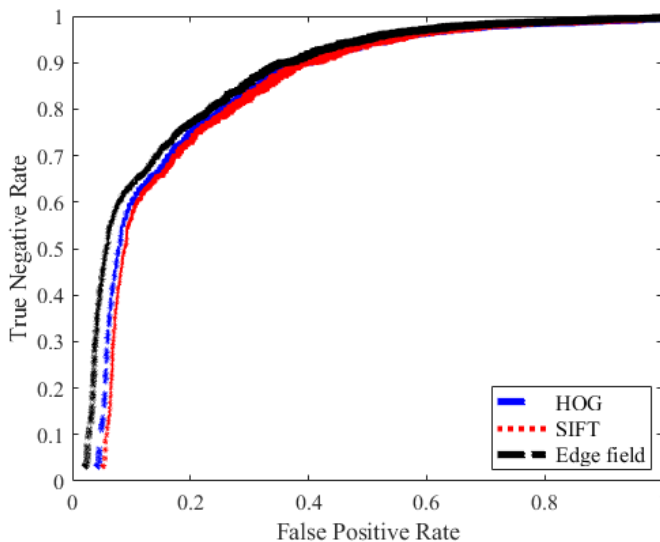


Figure 6. The ROC curve comparison



video, the initial particles are randomly sampled where the response of the detector is large. For the subsequent frames in the video, the initial particles are randomly sampled according to the estimation results of the previous frame. The posture detection results are reported in terms of confusion matrix which is shown in Figure 6.

From Figure 7, it can be found that the proposed posture recognition method can achieve 88.42% for all postures.

5. CONCLUSIONS

Aiming at the problem of human pose estimation in monocular images and videos, this paper adopts the component-based method, and uses the tree model containing 10 components to represent the human body. To prepare the observation model of each component, this paper proposes a rotation invariant feature of edge force field, and learns boosting classifier based on this feature. In addition, based on the idea of particle filter, this paper uses a group of particles in the state space to simulate the state distribution of components, and uses the belief propagation algorithm to calculate the marginal distribution of each component on this group of particles and update the particle position.

Figure 7. The confusion matrix of posture detection

	no action	walking	running	dribbling	jumping	shooting	passing	catching	recognitio n rate
no action	442	21	9	1	0	0	4	3	92.08%
walking	66	505	32	5	1	2	13	17	78.78%
running	2	2	483	9	6	3	37	23	85.49%
dribbling	5	1	3	450	8	3	17	29	87.21%
jumping	3	2	8	43	510	9	6	7	86.73%
shooting	10	1	2	19	11	519	3	4	91.21%
passing	2	6	1	7	1	1	458	8	94.63%
catching	10	2	2	6	3	3	2	449	94.13%
recognitio n rate	81.85%	93.52%	89.44%	83.33%	94.44%	96.11%	84.81%	83.15%	88.42%

The calculation of edge force field features is simple and expressive, which can be used to calculate the observation model of components quickly and accurately. The belief propagation algorithm based on particle sampling uses particle simulation state distribution to reduce the computational burden. Experiments show that the accuracy of this method is higher than that of previous algorithms. Our posture estimation method will be further used to develop a real application system for assisting daily teaching.

ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

FUNDING AGENCY

Publisher has waived the Open Access publishing fee.

REFERENCES

- Batool, M., Jalal, A., & Kim, K. (2019, August). Sensors Technologies for Human Activity Analysis Based on SVM Optimized by PSO Algorithm. In *2019 International Conference on Applied and Engineering Mathematics (ICAEM)* (pp. 145-150). IEEE.
- Bui, M. Q., Duong, V. H., Tai, T. C., & Wang, J. C. (2018, October). Depth Human Action Recognition Based on Convolution Neural Networks and Principal Component Analysis. In *2018 25th IEEE International Conference on Image Processing (ICIP)* (pp. 1543-1547). IEEE.
- Cai, C., Yang, S., Yan, P., Tian, J., Du, L., & Yang, X. (2019, August). Real-Time Human-Posture Recognition for Human-Drone Interaction Using Monocular Vision. In *International Conference on Intelligent Robotics and Applications* (pp. 203-216). Springer.
- Chevtchenko, S. F., Vale, R. F., Macario, V., & Cordeiro, F. R. (2018). A convolutional neural network with feature fusion for real-time hand posture recognition. *Applied Soft Computing*, 73, 748–766.
- Fan, J., Bi, S., Wang, G., Zhang, L., & Sun, S. (2021). Sensor Fusion Basketball Shooting Posture Recognition System Based on CNN. *Journal of Sensors*.
- Friston, K. J., Parr, T., & de Vries, B. (2017). The graphical brain: Belief propagation and active inference. *Network Neuroscience*, 1(4), 381–414.
- Grosse-Puppenthal, T., Holz, C., Cohn, G., Wimmer, R., Bechtold, O., Hodges, S., & Smith, J. R. et al. (2017, May). Finding common ground: A survey of capacitive sensing in human-computer interaction. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 3293-3315).
- Jalal, A., Akhtar, I., & Kim, K. (2020). Human Posture Estimation and Sustainable Events Classification via Pseudo-2D Stick Model and K-ary Tree Hashing. *Sustainability*, 12(23), 9814.
- Kumaran, S. K., Dogra, D. P., & Roy, P. P. (2019). *Anomaly detection in road traffic using visual surveillance: A survey*. arXiv preprint arXiv:1901.08292.
- Lahiani, H., Kherallah, M., & Neji, M. (2017, December). Hand gesture recognition system based on Local Binary Pattern approach for mobile devices. In *International Conference on Intelligent Systems Design and Applications* (pp. 180-190). Springer.
- Li, B., Bai, B., & Han, C. (2020). Upper body motion recognition based on key frame and random forest regression. *Multimedia Tools and Applications*, 79(7), 5197–5212.
- Li, K., & Huang, Q. (2019). Cross-pose face recognition by integrating regression iteration and interactive subspace. *EURASIP Journal on Wireless Communications and Networking*, 2019(1), 1–8.
- Mircoli, A., & Cimini, G. (2018, September). Automatic Extraction of Affective Metadata from Videos Through Emotion Recognition Algorithms. In *European Conference on Advances in Databases and Information Systems* (pp. 191-202). Springer.
- Murahari, V. S., & Plötz, T. (2018, October). On attention models for human activity recognition. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers* (pp. 100-103). Academic Press.
- Oudah, M., Al-Naji, A., & Chahl, J. (2020). Hand gesture recognition based on computer vision: a review of techniques. *Journal of Imaging*, 6(8), 73.
- Qi, J., Xu, K., & Ding, X. (2021). Approach to hand posture recognition based on hand shape features for human-robot interaction. *Complex & Intelligent Systems*, 1-18.
- Ren, H., & Wang, X. (2021). Application of wearable inertial sensor in optimization of basketball player's human motion tracking method. *Journal of Ambient Intelligence and Humanized Computing*, 1–15.
- Sharma, L., & Lohan, N. (2019). Performance analysis of moving object detection using BGS techniques in visual surveillance. *International Journal of Spatio-Temporal Data Science*, 1(1), 22–53.
- Tomić, I., & McCann, J. A. (2017). A survey of potential security issues in existing wireless sensor network protocols. *IEEE Internet of Things Journal*, 4(6), 1910–1923.

Wang, B., Gu, X., Ma, L., & Yan, S. (2017). Temperature error correction based on BP neural network in meteorological wireless sensor network. *International Journal of Sensor Networks*, 23(4), 265–278.

Wang, Z., Shi, X., Wang, J., Gao, F., Li, J., Guo, M., & Qiu, S. et al. (2019, October). Swimming Motion Analysis and Posture Recognition Based on Wearable Inertial Sensors. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)* (pp. 3371-3376). IEEE.

Xu, W. (2019). Toward human-centered AI: A perspective from human-computer interaction. *Interaction*, 26(4), 42–46.

Zhang, L. (2018). Research on human body movement posture based on inertial sensor. *International Journal Bioautomation*, 22(2), 179.